An extension of chaotic probability models to real-valued variables

P.I.Fierens
Department of Physics and Mathematics,
Instituto Tecnológico de Buenos Aires (ITBA),
Argentina
pfierens@itba.edu.ar

Abstract

Previous works ([5][6][8])) have presented a frequentist interpretation of sets of measures as probabilistic models which have denominated *chaotic models*. Those models, however, dealt only with sets of probability measures on finite algebras, that is, probability measures which can be related to variables with a finite number of possible values. In this paper, an extension of chaotic models is proposed in order to deal with the more general case of real-valued variables.

Keywords. Imprecise probabilities, foundations of probability, chaotic probability models, frequentist interpretation.

1 Introduction

In a series of papers ([5][6]), we presented the first steps towards a frequentist interpretation of sets of measures as probability models, which we have called *chaotic probability models* in order to distinguish them from other plausible interpretations. This work was coherently presented in [4] and extended by Rêgo and Fine in [8]. In our previous work, we presented chaotic models as simply sets of probability measures whose domain is a finite set of events. In this sense, we may associate chaotic probability models to discrete "random"[1] variables with finite range (e.g., the outcome of the flipping of a coin or the tossing of a die). In this paper, we present a simple approach to the extension of chaotic probability models to real-valued variables (e.g., tomorrow's minimum temperature).

The paper is organized as follows. Section 2 presents some concepts of the previous work which are needed for this paper. In Section 3, we provide the basic motivation behind the model which is described in Section 4. In the latter Section, we also show that such a model is plausible. Section 5 is devoted to present extensions for this framework of the concepts of visibility and temporal homogeneity defined in previous works. Finally, in Section 6 we discuss the results presented in this paper and suggest future lines of work.

[1] We use quotation marks to denote the difference between these *chaotic* variables and the usual understanding of random variables.

2 Variables with finite range

We need to recall the interpretation of chaotic probability models for variables with finite range ([6][4][8]).

2.1 An Instrumental Description of the Model

The instrumental (that is, without commitment to reality) description of chaotic probability models presented in earlier works is basically preserved in this paper. Let $\mathbf{X}$ be a finite sample space. We denote by $\mathbf{X}^*$ the set of all finite sequences of elements taken in $\mathbf{X}$. A particular sequence of n samples from $\mathbf{X}$ is denoted by $x^n = \{x_1, x_2, \cdots, x_n\}$. $\mathbf{P}$ denotes the set of all measures on the power set of $\mathbf{X}$. A chaotic probability model $\mathbf{M}$ is a subset of $\mathbf{P}$ and models the "marginals" of some process generating sequences in $\mathbf{X}^*$.

Given any $n \in \mathbb{N}$, consider the generation of a sequence x^n of length n by the following algorithm[2]:

`FOR k = 1 TO k = n`

1. `Choose` $\nu = F(x^{k-1}) \in \mathbf{M}$.
2. `Generate` x_k `according to` ν.

where $F : \mathbf{X}^* \to \mathbf{M}$ is a function corresponding to the decisions causally made by the algorithm at each step. Let $\nu_k = F(x^{k-1})$. For any $k \leq n$, F determines the probability distribution of the *potential* kth outcome X_k of the sequence,

$$(\forall \mathbf{A} \subseteq \mathbf{X}) \;\; P\left(X_k \in \mathbf{A} | X^{k-1} = x^{k-1}\right) = \nu_k(X_k \in \mathbf{A}).$$

[2] We denote the empty string by x^0.

The probability of a particular realization x^n of a sequence of random variables X^n is given by

$$P\left(X_1 = x_1, \ldots, X_n = x_n\right) = \prod_{k=1}^{n} \nu_k(X_k = x_k).$$

We denote by $\mathbf{M}^*$ the family of all such process measures P, one for each possible function F. From the analysis of data, we do not expect in general to be able to pinpoint a single $P \in \mathbf{M}^*$ or even a small subset of $\mathbf{M}^*$, what we call a **fine-grained picture** of the source. On the contrary, we expect our knowable **operational quantities to be (large) subsets of $\mathbf{M}^*$** which provide an appropriate **coarse-grained** description of the source.

2.2 Data analysis and estimation

We begin the study of a sequence $x^n \in \mathbf{X}^*$ by analyzing it into several subsequences. These subsequences are selected by rules that satisfy the following

Definition 1. *A computable function $\psi : \mathbf{X}^* \to \{0,1\}$ is a* **causal subsequence selection rule** *(also known as a Church place selection rule) if for any $x^n \in \mathbf{X}^*$, x_k is the j-th term in the generated subsequence $x^{\psi,n}$, of length $\lambda_{\psi,n}$, whenever*

$$\psi(x^{k-1}) = 1,\ \sum_{i=1}^{k} \psi(x^{i-1}) = j,\ \lambda_{\psi,n} = \sum_{k=1}^{n} \psi(x^{k-1}).$$

Let $\Psi = \{\psi_\alpha\}$ be a set of causal subsequence selection rules. For each $\psi \in \Psi$, we study the behavior of the relative frequency of marginal events along the chosen subsequence. That is, given x^n and a selection rule $\psi \in \Psi$ we determine the **frequentist empirical (relative frequency) measure** $\bar{\mu}_{\psi,n}$ along the subsequence $x^{\psi,n}$ through

$$(\forall \mathbf{A} \subset \mathbf{X})\ \bar{\mu}_{\psi,n}(\mathbf{A}) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} I_{\mathbf{A}}(x_k)\psi(x^{k-1}),$$

where $I_{\mathbf{A}}(\cdot)$ is the indicator function of the event $\mathbf{A}$.

A family of subsequence selection rules Ψ is key to our understanding of a chaotic probability model as given by a set of measures $\mathbf{M}$. It has been proved that:

- so long as we restrict to a family of causal selection rules of moderate size, we can with high probability avoid extracting arbitrary patterns through some of the selected subsequences and instead exhibit the patterns that have inductive validity (see [6] and [4]).
- chaotic probability models $\mathbf{M}$ can be estimated from the empirical relative frequency measures if the appropriate family of subsequence selection rules is chosen (see [6] and [4]). Rêgo and Fine [8] showed how to choose a universal family of place selection rules to make the model visible.
- the visibility (possibility of estimation) of a chaotic probability model $\mathbf{M}$ depends strongly on the choice of the subsequence selection family, i.e., there are cases where $\mathbf{M}$ can be estimated by a family Ψ_0 while another family Ψ_1 only "sees" one measure in $ch(\mathbf{M})$ (see [6] and [4]).

3 Motivation

In what follows, we shall assume that $(\mathbf{X}, \mathcal{X})$ is a measurable space and that $\mathbf{P}$ is the set of all probability measures on $\mathcal{X}$. A chaotic probability model is represented mainly by a set $\mathbf{M} \subset \mathbf{P}$.

The instrumental description of chaotic probability models summarized in Section 2.1 can be extended to variables with infinite (even uncountable) range without changes. Therefore, the problem of the extension of chaotic probability models to more general spaces lies on the task of making such models "visible" (in an intuitive sense) when they are represented as in Section 2.1.

One possibility is to allow, as in the finite case, for the estimation of the measures in $\mathbf{M}$ by means of the empirical relative frequencies:

$$\bar{\mu}_{\psi,n}(\mathbf{A}) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} I_{\mathbf{A}}(x_k)\psi(x^{k-1}).$$

The difficulty then becomes the choice of the sets $\mathbf{A} \subset \mathbf{X}$ that should be used. For, in general, it is impossible to compute $\bar{\mu}_{\psi,n}(\mathbf{A})$ for all $\mathbf{A}$ in a σ-field. Furthermore, it may make no sense at all to try to assess such a fine-grained model.

We may also charge the statistician with the responsibility of choosing a collection of subsets $\mathbf{A} \subset \mathbf{X}$ adequate for the problem at hand. If we follow this path, we may as well allow for greater generality by letting the practitioner to choose a suitable finite family $\mathbf{F}$ of real-valued bounded measurable *test functions* $f : \mathbf{X} \to \mathbb{R}$ and proceeding to the estimate by means of the empirical relative frequencies

$$\bar{\mu}_{\psi,n}(f) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} f(x_k)\psi(x^{k-1}).$$

We may conceive these functions as those which are of actual interest for the problem at hand. We may also understand the family of functions $\mathbf{F}$ together with the family of subsequence selection rules Ψ as a representation of the discernment power of the observer

or, at least, of the coarse-grainedness appropriate for the model. In particular note that if we restrict ourselves to bounded test functions, the family $\mathbf{F}$ may contain indicator functions of the type $I_{\mathbf{A}}$, $\mathbf{A} \subset \mathbf{X}$.

From a technical viewpoint, the trick is simple: we substitute the *finite* algebra of events related to a discrete variable by a *finite* set of test functions applied to a real variable. With this idea in mind, all previous results (e.g., those in [6]) can be easily extended, as it is shown in Section 5.

3.1 Test functions as gambles

From a behavioral stand, we may also consider $\mathbf{F}$ as a collection of gambles in the sense of Peter Walley:

> *"A gamble is a bounded real-valued function on* Ω [the sample space] *which is interpreted as a reward."* ([11], Chapter 2)

Therefore, from this point of view, the estimates $\bar{\mu}_{\psi,n}(f)$ can be understood as estimates of the set of linear previsions dominating a coherent lower prevision based on the gambles $\mathbf{F}$ (see [11], especially Chapters 2 and 3).

There is, however, a key difference with some of the work of Peter Walley in [11] in the sense that we are not interested in pursuing an equivalent to the natural extension in our framework. We assume that a finite set of test functions (or gambles) $\mathbf{F}$ is enough for the purposes of a given problem. In other words, we do not feel compelled to make probabilistic assessment over anything else than $\mathbf{F}$.

If we take the discussion in the last paragraph one step further, we may allow to specify a chaotic probability model by, not a precise set of probability measures $\mathbf{M}$, but by a collection of "previsions" $\mathbf{P}_r$ defined as

$$(\forall f \in \mathbf{F})\, \mathbf{P}_r(f) \subset \mathbb{R},\ \mathbf{P}_r = \{\mathbf{P}_r(f) :\ f \in \mathbf{F}\},$$

and where, intuitively,

$$\mathbf{P}_r(f) = \{\mu(f) :\ \mu \in \mathbf{M}\}$$

for some unspecified set $\mathbf{M}$.

There is another difference with the work of Peter Walley: we consider only countably additive linear previsions $\mathbf{P}$. Theorem 3.6.1 of Walley [11] shows that there is a one to one correspondence between the coherent lower previsions on the domain of the bounded gambles and the non-empty weak*-compact convex subsets $\mathbf{M}$ of $\mathbf{P}_l$, the set of all linear previsions. Although $\mathbf{P}_l$ is the weak*-closure of $\mathbf{P}$, the latter is strictly contained in the former if the sample space $\mathbf{X}$ is infinite (see [11], Appendix D and Section 3.6.8).

Since we do not desire to follow closely any behavioral interpretation and in order to avoid confusion, we continue to refer to the elements of $\mathbf{F}$ as test functions rather than gambles.

3.1.1 Gambling with nature

Shafer and Vovk ([10]) present a game-theoretic interpretation of probabilistic reasoning where there are three players:

1. `Nature`: It determines what may happen in the world, that is, the outcomes of a given game.
2. `Modeler`: `Modeler` suggests a theory about how `Nature` behaves. Based on this theory, `Modeler` proposes a game to Skeptic.
3. `Skeptic`: `Skeptic` tries to show that `Modeler`'s theory is wrong by betting on the proposed game. If `Modeler`'s theory is "correct", `Skeptic` should not be able to make a high profit (with "not-too-low probability").

It is easy to see a relationship between the work presented here and the theory of Shafer and Vovk. `Modeler` has an instrumental understanding of how `Nature` works (see Section 2.1): There is a certain set of measures $\mathbf{M} \subset \mathbf{P}$ such that

`FOR k = 1 TO k = n`

1. `Nature chooses` $\nu \in \mathbf{M}$ `based on` x^{k-1}.
2. `Nature generates` x_k `according to` ν.

Note that `Nature` is *causal, but not necessarily markovian.* Also note that `Modeler` does not need to know the set $\mathbf{M}$, but `Modeler` does know $\{\mu(f)\}_{\mu \in \mathbf{M}}$ for all $f \in \mathbf{F}$. Based on this understanding of `Nature`'s behavior, `Modeler` proposes any of the following gambles to `Skeptic`:

`Skeptic's initial capital is` $K_0 = 0$.

`Skeptic chooses` $f \in \mathbf{F}$.

`FOR k = 1 TO k = n`

1. `Skeptic chooses` $\theta_k \in \{0, 1\}$ `based on` x^{k-1}.
2. `Nature generates` x_k.
3. `Skeptic's capital is now`

$$K_k = K_{k-1} + \theta_k \left(f(x_k) - \sup_{\mu \in \mathbf{M}} \mu(f) \right).$$

Note that `Skeptic`'s sequence of bets $\{\theta_k\}$ can be associated with a causal subsequence selection rule ψ (see Def. 1), if `Skeptic` can only use computable strategies. Note also that, if `Skeptic` needs to keep track of his capital, he must be able to compute his earnings. One necessary (but not sufficient) requirement for this is that the test functions be computable in a reasonable sense (see Section 3.2). We shall come back to this game in Section 4.

3.2 Computability of test functions and place selection rules

One reasonable restriction on test functions is to ask them to be computable, i.e., that their value can be calculated. Also, place selection rules must be able to output values in $\{0, 1\}$ when having tuples of real values as inputs. The problem then becomes to find a reasonable definition of computable real-valued functions with real-valued input variables. To our knowledge, there are mainly two broad approaches to such a definition in the area of *computational analysis* (see, e.g., [2] [1]). On one hand, there are the traditional approach and its many variations and extensions which are based, loosely speaking, on the following ideas:

- Given a finite alphabet, say $\mathbf{A}$, and an adequate program M, a *description* of an element y of some space $\mathbf{Y}$ is a finite string $\underline{a} = a_1 a_2 \cdots a_k$, $a_i \in \mathbf{A}$, such that $y = M(\underline{a})$.
- It may be the case that not all the elements of some space have a description. However, an element y is considered to be computable if there is an *approximating sequence of descriptions* $\{\underline{a}_i\}$, i.e., the strings $\underline{a}_i$ are such that the outputs $y_i = M(\underline{a}_i)$ get increasingly closer to y.
- A function $f : \mathbf{Y} \to \mathbf{Z}$ is computable if for each computable number $y \in \mathbf{Y}$, with approximating sequence $\{\underline{a}_i\}$, there is a program P such that $\{P(\underline{a}_i)\}$ is an approximating sequence of descriptions for $f(y) \in \mathbf{Z}$.

This approach models well scientific computations. Moreover, most "calculator" functions (polynomials, $\log(x)$, $\sqrt{x}$, etc.) are computable under this approach.

On the other hand, there is the Blum-Shub-Smale (BSS) approach which is based on computing machines that can deal with elements of any field R (e.g, $R = \mathbb{R}$) and that are allowed to perform the field operations ($+$, $-$, $\times$ and $\%$) on R and can branch on comparisons ($<$, $>$, $\leq$) between elements of R if it is ordered. The fact that the BSS approach is very useful in numerical modelling should not come as a surprise.

Since we are focused on calculations that can be made on any personal computer, we shall take the first approach to computability of real-valued functions.

3.2.1 Computable functions of real variable

The material in this section is taken from Weihrauch [12] (see also [7]). There are other approaches which are equivalent and quite powerful, for example, that based on domain theory (see, e.g., Edalat [3]), but less intuitive.

Let $\mathbf{A}$ be any finite alphabet. The finite strings of elements of $\mathbf{A}$ will be denoted by $\mathbf{A}^*$ and the infinite sequences of elements of $\mathbf{A}$ will be denoted by $\mathbf{A}^\infty$.

Definition 2. (Computability by Type 2 machines)

1. *A* **Type 2 machine** *M is defined by two components:*
   (a) *a Turing machine with k one-way input tapes ($k \geq 0$), a single one-way output tape and finitely many work tapes,*
   (b) *a type specification $(\mathbf{Y}_1, \cdots, \mathbf{Y}_k, \mathbf{Y}_0)$ with $\{\mathbf{Y}_0, \cdots, \mathbf{Y}_k\} \subseteq \{\mathbf{A}^*, \mathbf{A}^\infty\}$.*
2. *The function $\rho_M : (\mathbf{Y}_1 \times \cdots \times \mathbf{Y}_k) \to \mathbf{Y}_0$ computed by the Type 2 machine M (the semantics of M) is defined as follows:*
   (a) *Case $\mathbf{Y}_0 = \mathbf{A}^*$ (finite output): $\rho_M(y_1, \cdots, y_k) = w$ iff M with input $(y_1, \cdots, y_k)$ halts with result w on the output tape.*
   (b) *Case $\mathbf{Y}_0 = \mathbf{A}^\infty$ (infinite output): $\rho_M(y_1, \cdots, y_k) = p$ iff M with input $(y_1, \cdots, y_k)$ computes forever writing the sequence p on the output tape.*
3. *We say that a function $\rho : (\mathbf{Y}_1 \times \cdots \times \mathbf{Y}_k) \to \mathbf{Y}_0$ is* **computable** *iff $\rho = \rho_M$ for some Type 2 machine M. A sequence y is a* **computable element of $\mathbf{Y}_0$** *iff the 0-place function $\rho : \{()\} \to \mathbf{Y}_0$ with $\rho() = y$ is computable.*

Type 2 machines can be considered as a certain kind of oracle Turing machines and computability with respect to them is entirely classical. In order to extend the concept of computability to functions, e.g., over the reals, we need the concept of a naming system. Indeed, objects like real numbers can be represented (named) by finite or infinite sequences of finite alphabets. For example, we can represent a real number

in $[0,1]$ by its (probably infinite) representation by a binary sequence. These ideas are formalized in the following

Definition 3. (Naming System. Reducibility)

1. *A* **notation** *of a set* $\mathbf{X}$ *is a surjective function* $\rho : \mathbf{A}^* \to \mathbf{X}$ *(naming by finite strings).*
2. *A* **representation** *of a set* $\mathbf{X}$ *is a surjective function* $\rho : \mathbf{A}^\infty \to \mathbf{X}$ *(naming by infinite sequences).*
3. *A* **naming system** *of a set* $\mathbf{X}$ *is a notation or a representation of* $\mathbf{X}$*.*
4. *For functions* $\gamma : \mathbf{Y} \to \mathbf{X}$ *and* $\gamma' : \mathbf{Y}' \to \mathbf{X}'$ *with* $\mathbf{Y}, \mathbf{Y}' \subseteq \{\mathbf{A}^*, \mathbf{A}^\infty\}$*, we call* γ **reducible** *to* γ'*,* $\gamma \preceq \gamma'$*, iff there exists a computable function* $\rho : \mathbf{Y} \to \mathbf{Y}'$ *such that* $(\forall y \in dom(\gamma))\ \gamma(y) = \gamma'(\rho(y))$*. We say that* γ *and* γ' *are* **equivalent***,* $\gamma \equiv \gamma'$*, iff* $\gamma \preceq \gamma'$ *and* $\gamma' \preceq \gamma$*.*

In order to clarify ideas, we present some common naming systems:

- *Binary representation of* $\mathbb{N}$: $\rho_{bin} : \{0,1\}^* \to \mathbb{N}$, $\rho_{bin}(a_0 a_1 \cdots a_k) = \sum_{i=0}^{k} a_i 2^i$.
- *Rational numbers*: $\rho_{\mathbb{Q}} : \{+,-\} \times \{0,1\}^* \times \{0,1\}^* \to \mathbb{Q}$, $\rho_{\mathbb{Q}}(s, b_n, b_d) = s\frac{\rho_{bin}(b_n)}{\rho_{bin}(b_d)}$.
- *Interval Representation of* $\mathbb{R}$: Let $\mathbf{S}_{\mathbb{Q}}$ be the set of all infinite sequences of triples (s,n,d) taken from $\{+,-\} \times \{0,1\}^* \times \{0,1\}^*$. Then define $\rho_{int} : \mathbf{S}_{\mathbb{Q}} \times \mathbf{S}_{\mathbb{Q}} \to \mathbb{R}$ by

$$\rho_{int}(a_0 a_1 a_2 \cdots, b_0 b_1 b_2 \cdots) = x \Leftrightarrow$$
$$\Leftrightarrow \lim_{n\to\infty} \rho_{\mathbb{Q}}(a_n) = \lim_{n\to\infty} \rho_{\mathbb{Q}}(b_n) = x$$

and

$$\rho_{\mathbb{Q}}(a_0) < \rho_{\mathbb{Q}}(a_1) < \cdots < x < $$
$$< \cdots < \rho_{\mathbb{Q}}(b_1) < \rho_{\mathbb{Q}}(b_0).$$

The latter naming system leads to the following

Definition 4. (Computable Real Numbers) $x \in \mathbb{R}$ *is computable if it is* ρ_{int}*-computable.*

The definition of naming systems leads to the extension of the definition of computable functions that we need for this paper:

Definition 5. (Relative Computability)

1. *For* $i = 0, 1, \cdots, k$*, let* $\gamma_i : \mathbf{Y}_i \to \mathbf{Z}_i$ *be naming systems. A function* $\delta : \mathbf{Z}_1 \times \cdots \times \mathbf{Z}_k \to \mathbf{Z}_0$ *is* $(\gamma_1, \cdots, \gamma_k, \gamma_0)$*-computable iff there is a Type 2-computable function (in the sense of Def. 2)* $\rho : \mathbf{Y}_1 \times \cdots \times \mathbf{Y}_k \to \mathbf{Y}_0$ *such that*

$$\delta(\gamma_1(y_1), \gamma_2(y_2), \cdots, \gamma_k(y_k)) = $$
$$= \gamma_0(\rho(y_1, y_2, \cdots, y_k)),$$

whenever $\delta(\gamma_1(y_1), \gamma_2(y_2), \cdots, \gamma_k(y_k))$ *exists.*

2. *We say that a real-valued function of a real variable is computable if it is* (ρ_{int}, ρ_{int})*-computable.*

One important consequence of the definition of computability is that all computable functions are continuous (see [12]).

We shall require all *admissible test functions* to be computable. Some examples of real-valued computable functions are: $+$, $-$, $\times$, $1/x$, exp, log, sin, cos, $\sqrt{}$, min, max, etc.

We shall also require *place selection rules* to be computable functions of tuples of real variables, in the sense of Definition 5, which take only values in $\{0,1\}$. In other words, we shall require of a place selection rule ψ to be $(\rho_{int}, \cdots, \rho_{int})$-computable, where ρ_{int} appears $k+1$ times, for each $k \geq 0$.

We shall also need the following

Definition 6. (Computable Probability Mass Function) *Let* $(\mathbf{X}, \mathcal{X})$ *be a measurable space, with* $\mathcal{X}$ *containing the singleton sets. Then, we say that a probability mass function on* $(\mathbf{X}, \mathcal{X})$ *is computable if each of the probability values is computable in the sense of Def. 4.*

4 Chaotic probability model

Let $\Psi = \{\psi_\alpha\}$ be a set of causal subsequence selection rules and $\mathbf{F} = \{f_\beta\}$ a collection of bounded real-valued test functions. For each $\psi \in \Psi$, we study the behavior of the relative frequency of (only) f_β along the chosen subsequence. That is, given x^n and a selection rule $\psi \in \Psi$ we determine the **frequentist empirical (relative frequency) measure** $\bar{\mu}_{\psi,n}$ along the subsequence $x^{\psi,n}$ through

$$(\forall f \in \mathbf{F})\ \bar{\mu}_{\psi,n}(f) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} f(x_k)\psi(x^{k-1}).$$

In a similar manner, for all such rules ψ, we define the **time average conditional measure** $\bar{\nu}_{\psi,n}$ $(\forall f \in \mathbf{F})$

$$\bar{\nu}_{\psi,n}(f) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} \mathrm{E}\left[f(X_k) | X^{x-1} = x^{k-1}\right] \psi(x^{k-1}).$$

Rewritten in terms of our instrumental understanding of the measure selection function F,

$$\bar{\nu}_{\psi,n}(f) = \frac{1}{\lambda_{\psi,n}} \sum_{k=1}^{n} \nu_k(f)\psi(x^{k-1}),$$

where $\nu_k = F(x^{k-1})$. Note that, since we assume F to be unknown, the time average conditional measure $\bar{\nu}_{\psi,n}$ is also unknown. Since we want to expose some of the structure of the chaotic probability model $\mathbf{M}$ by means of the rules in Ψ, we are interested in how good $\bar{\mu}_{\psi,n}$ is as an estimator of $\bar{\nu}_{\psi,n}$.

Define the metric $d_{\mathbf{F}}$ on $\mathbf{P}$ by

$$d_{\mathbf{F}}(\nu,\mu) = \max_{f\in\mathbf{F}} |\mu(f) - \nu(f)|, \ (\forall \mu,\nu \in \mathbf{P}).$$

We call **F-causally faithful** a set of rules Ψ such that any $\psi \in \Psi$ yields a small value of $d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \bar{\mu}_{\psi,n})$ with high probability. The existence of such a set of rules is stated by

Theorem 1. *Let $m \leq n$ and fix Ψ and $\mathbf{F}$ of finite cardinality, denoted by $\|\Psi\|$ and $\|\mathbf{F}\|$ respectively. Then* $(\forall P \in \mathbf{M}^*)$

$$P\left(\max_{\psi\in\Psi}\{d_{\mathbf{F}}(\bar{\mu}_{\psi,n},\bar{\nu}_{\psi,n}) : \ \lambda_{\psi,n} \geq m\} \geq \varepsilon\right) \leq \\ \leq 2\|\mathbf{F}\|\|\Psi\|\mathrm{e}^{-\frac{\varepsilon^2 m^2}{8\beta^2 n}},$$

where

$$\beta = \max_{f\in\mathbf{F}} \sup_{x\in\mathbf{X}} |f(x)|.$$

The proof of the theorem is completely analog to that of Theorem 1 in [6] (see also the appendix to Chapter 4 in [4]). The consequence of this theorem is that, as long as we restrict to small-sized families of causal selection rules we can with high probability avoid extracting arbitrary patterns through some of the selected subsequences.

Recall the game in Section 3.1.1 proposed by `Modeler` to `Skeptic`. If `Modeler` is right, the probability that `Skeptic` becomes rich is very low. This is exactly what the following result shows.

Lemma 1. *Consider the game played by* `Modeler` *and* `Skeptic`. *Then* $(\forall \varepsilon > 0)$ $(\forall m \leq n)$

$$P(K_n \geq m\varepsilon) \leq 2\mathrm{e}^{-\frac{\varepsilon^2 m^2}{8\beta^2 n}},$$

where

$$\beta = \max_{f\in\mathbf{F}} \sup_{x\in\mathbf{X}} |f(x)|.$$

The proof of this lemma follows along the same lines as the proof of Theorem 1.

4.1 Collection of expected values as a model

In Section 3.1, we suggested the idea of taking the collection of expected values as the actual model, defining implicitly the set of probability measures $\mathbf{M}$. The following Lemma shows that $\mathbf{M}$ defined in this way has a particularly simple structure.

Lemma 2. *Let $(\mathbf{X},\mathcal{X})$ be a measurable space and $\mathbf{P}$ the set of all probability measures on it. Assume that $\mathcal{X}$ contains the singletons. Let $\mathbf{F} = \{f_1,\cdots,f_N\}$ be a finite collection of real-valued bounded functions. Let the set*

$$\mathbf{P}_r \subset \left[\inf_{x\in\mathbf{X}} f_1(x), \sup_{x\in\mathbf{X}} f_1(x)\right] \times \cdots \\ \cdots \times \left[\inf_{x\in\mathbf{X}} f_N(x), \sup_{x\in\mathbf{X}} f_N(x)\right] \subset \mathbb{R}^{\|\mathbf{F}\|}$$

be given. Define a set of measures by

$$\mathbf{M}_{\mathbf{P}_r} = \{\mu \in \mathbf{P} : \ (\mu(f_1),\cdots,\mu(f_N)) \in \mathbf{P}_r\}.$$

Then, the measures in $\mathbf{M}_{\mathbf{P}_r}$ are ε-indistinguishable from measures with finite support in the sense that for each $\varepsilon > 0$ there are points $x_1,\cdots,x_{L(\varepsilon)}$ in $\mathbf{X}$ such that $(\forall \mu \in \mathbf{M}_{\mathbf{P}_r})(\exists \nu \in \mathbf{M}_{\mathbf{P}_r})$ such that

$$d_{\mathbf{F}}(\mu,\nu) \leq \varepsilon, \ \textit{and} \ \sum_{i=1}^{L(\varepsilon)} \nu(\{x_i\}) = 1.$$

In other words, Lemma 2 tells us that, as long as we restrict ourselves to a *finite* set of test functions, there is no substantial difference (what concerns the test functions) between the behavior of a given *chaotic* real variable and that of a particular *chaotic* discrete variable with finite range. This fact not only opens up the door to the reuse of previous results which were originally conceived for discrete variables, but it also shows the way in which chaotic real variables can be simulated. Indeed, the simulation of chaotic real variable is not different from that of an adequate chaotic discrete variable according to Lemma 2, and the simulation of the latter type of variables was already explained in [6] (see also the proof of Theorem 3 in the Appendix).

Hence, using a collection of expected values of a finite set of test functions $\mathbf{P}_r$ as a model, gives us only a coarse-grained, blurred view of how a real variable behaves. This model may be as precise as we are capable of (or willing to) build it. However, the model is so fuzzy, our view so blurred, that we cannot distinguish with certainty whether we observe a real-valued variable or just a simple discrete variable which takes only a few values. By the way, this should not be

very surprising for, that who observes a finite number of outcomes of a uniformly distributed random variable in $[0,1]$, how can he be certain that he was dealing with a real random variable or just a complex discrete random variable.

5 Visibility and Temporal Homogeneity

In this section, we present extensions to those concepts of visibility and temporal homogeneity which were defined in [6]. The proofs of the results that follow are also analog to the proofs of the results in [6] thanks to the finiteness of the set of bounded test functions $\mathbf{F}$ and Lemma 2.

The possibility of exposing all of $\mathbf{M}$ by means of the rules in Ψ is expressed in the following

Definition 7. (**Visibility**)

(a) $\mathbf{M}$ *is* **made F-visible** (Ψ,θ,δ,m,n) *by* $P\in\mathbf{M}^*$ *if*

$$P\left(\bigcap_{\mu\in\mathbf{M}}\bigcup_{\psi\in\Psi}\mathbf{C}_\psi\right)\geq 1-\delta,$$

where

$$\mathbf{C}_\psi=\{X^n:\lambda_{\psi,n}(X^n)\geq m, d_{\mathbf{F}}(\bar{\mu}_{\psi,n},\mu)\leq\theta\}.$$

(b) *A subset* $\mathbf{M}'$ *of* $\mathbf{M}^*$ **renders M uniformly F-visible** (Ψ,θ,δ,m,n) *if* $\mathbf{M}$ *is made* $\mathbf{F}$*-visible* (Ψ,θ,δ,m,n) *by each* $P\in\mathbf{M}'$*. The maximal such subset is denoted* $\mathbf{M}_V(\Psi)$ *and* $\mathbf{M}_V(\Psi)$ *may be empty.*

The non-triviality of Definition 7**(a)**, and, hence, of Definition 7**(b)**, is asserted in

Theorem 2. *Let* $\mathbf{M}$ *be a set of probability measures and* $\mathbf{F}$ *a finite family of real-valued bounded functions on* $\mathbf{X}$*. Given* $0<2\varepsilon<\theta$*, for large* n*, there exists a process measure* P *and a family* Ψ *of size* N_ε *such that* $\mathbf{M}$ *is made* $\mathbf{F}$*-visible* (Ψ,θ,δ,m,n) *by* P *with*

$$\delta=2(\|\mathbf{F}\|+1)N_\varepsilon \mathrm{e}^{-\frac{(\theta-2\varepsilon)^2m^2}{8\beta^2n}},$$

where

$$\beta=\max_{f\in\mathbf{F}}\sup_{x\in\mathbf{X}}|f(x)|,$$

$$N_\varepsilon\leq\left\lceil\frac{2\beta}{\varepsilon}\right\rceil^{\|\mathbf{F}\|}.$$

The fact that not every set of rules Ψ can expose all of $\mathbf{M}$ is expressed by the concept of temporal homogeneity defined as follows.

Definition 8. (**Temporal Homogeneity**)

(a) $P\in\mathbf{M}^*$ *is* **F-temporally homogeneous** (Ψ,θ,δ,m,n) *if*

$$P\left(\Delta_\Psi\leq\theta\right)\geq 1-\delta,$$

where

$$\Delta_\Psi=$$
$$=\max_{\psi_1,\psi_2\in\Psi}\left\{d_{\mathbf{F}}(\bar{\mu}_{\psi_1,n},\bar{\mu}_{\psi_2,n}):\lambda_{\psi_1,n},\lambda_{\psi_2,n}\geq m\right\}.$$

(b) *A subset* $\mathbf{M}'$ *of the set of all possible process measures* $\mathbf{M}^*$ *is* **uniformly F-temporally homogeneous** (Ψ,θ,δ,m,n) *if each of the elements of* $\mathbf{M}'$ *is temporally homogeneous* (Ψ,θ,δ,m,n)*. The maximal such subset is denoted* $\mathbf{M}_T(\Psi)$*.*

As it was the case with chaotic variables with finite range, a model $\mathbf{M}$ may be visible under a certain family of subsequence selection rules and temporal homogeneous under another, as the following result shows.

Theorem 3. *Let* $\mathbf{F}$*,* $\mathbf{P}_r$ *and* $\mathbf{M}_{\mathbf{P}_r}$ *be as in Lemma 2. Let* $\varepsilon>\frac{\beta}{m}$*, where*

$$\beta=\max_{f\in\mathbf{F}}\sup_{x\in\mathbf{X}}|f(x)|.$$

Let Ψ_0 *be a set of (causal deterministic) place selection rules. Then, there are a process measure* P *and a family* Ψ_1 *such that, for large enough* n*,* P *will both render* $\mathbf{M}_{\mathbf{P}_r}$ $\mathbf{F}$*-visible* $(\Psi_1,3\varepsilon,\delta,m,n)$ *and ensure* $\mathbf{F}$*-temporal homogeneity* $(\Psi_0,6\varepsilon,\delta,m,n)$ *with*

$$\delta=2\|\mathbf{F}\|\max\{\|\Psi_0\|,\|\Psi_1\|\}\,\mathrm{e}^{-\frac{\varepsilon^2m^2}{8\beta^2n}}.$$

Although the proof of this theorem is very similar to that of Theorem 4 in [6], we include a sketch in the appendix because it shows clearly how the concepts of computability of real-valued functions, the finiteness of the set of test functions $\mathbf{F}$ and Lemma 2 are applied in order to reuse previous results under the current framework.

6 Conclusions and future work

The extension of chaotic probability models proposed in this paper does not carry in itself any technical novelties with respect to previous works, except perhaps for Lemma 2. Although this may seem disappointing, we believe it is the best feature of the current presentation, i.e., that it allows a smooth and simple extension of chaotic models to real-valued variables.

Besides extending previous works on chaotic models, a different viewpoint on them is offered in Section

4.1, where we suggest to get rid of the set of measures and work directly with the assessment of "expected" values of the test functions. Although this idea is not novel in itself, it is in the framework of chaotic models.

The relation between gambles and test functions sketched in Section 3.1 may allow to those pursuing behavioral interpretations of probability to deal with chaotic models without any sense of guilt.

Lemma 2 shows that the finiteness of our discernment is implicitly embedded in the finite number of test functions.

There are several matters which were left out of this paper. For example, it is easy to see that the same ideas can be applied to tuples of variables. Then, the question becomes what the relation is between chaotic models on tuples of variables and the "marginal" chaotic models and how independence can be characterized. The problem of marginalizing chaotic models on tuples is difficult because the corresponding test functions must also be marginalized.

Acknowledgements

This paper would not have existed had it not been for the encouragement from T.L. Fine and Leandro Rêgo. The author is most grateful to T.L. Fine for he found a mistake in an earlier version of Lemma 2. The author would also like to thank the unknown reviewers for their helpful and useful comments.

This work was partially supported by anonymous contributors through the project "Prevention and early detection of forest fires by means of sensor networks" which is being developed at the Instituto Tecnológico de Buenos Aires (ITBA).

Appendix A: Proof of Lemma 2

In order to prove Lemma 2, we need the following preliminary result.

Proposition 1. *Let $\varepsilon > 0$ be given and let $\beta > 0$ be defined as*

$$\beta = \max_{f \in \mathbf{F}} \sup_{x \in \mathbf{X}} \|f(x)\|.$$

Then, there is a finite set $\mathbf{M}_\varepsilon = \{\nu_1, \nu_2, \cdots, \nu_{N_\varepsilon}\} \subset \mathbf{M}_{\mathbf{P}_r}$ such that

$$N_\varepsilon \leq \left\lceil \frac{2\beta}{\varepsilon} \right\rceil^{\|\mathbf{F}\|},$$

and

$$\sup_{\mu \in \mathbf{M}} \min_{1 \leq i \leq N_\varepsilon} d_{\mathbf{F}}(\mu, \nu_i) \leq \varepsilon.$$

Proof. Let $N = \|\mathbf{F}\|$ and consider the set in $\mathbb{R}^N$

$$\mathbf{A} = \{(\mu(f_1), \cdots, \mu(f_N)) : \mu \in \mathbf{M}_{\mathbf{P}_r}\}.$$

Then, it is clear that $\mathbf{A}$ is included in the closed hypercube $[-\beta, +\beta]^N$. Moreover, this hypercube can be covered by a set of $\left\lceil \frac{2\beta}{\varepsilon} \right\rceil^N$ smaller hypercubes of side ε. □

We also need the following result from Rudin [9] (see Lemma after Theorem 3.25 in Rudin [9], page 73).

Lemma 3. *If y lies in the convex hull of a set $\mathbf{E} \subset \mathbb{R}^N$, then y lies in the convex hull of a subset of $\mathbf{E}$ which contains at most $N + 1$ points.*

Now, we are ready for the proof of Lemma 2:

Proof. Let $\underline{f}_x = (f_1(x), \cdots, f_N(x))$ for all $x \in \mathbf{X}$, and $\underline{\mu} = (\mu(f_1), \cdots, \mu(f_N))$ for all $\mu \in \mathbf{M}_{\mathbf{P}_r}$. Consider the following set:

$$\mathbf{E} = \left\{ \underline{f}_x : x \in \mathbf{X} \right\} \subset \mathbb{R}^N.$$

It is clear that $\mathbf{P}_r \subseteq ch(\mathbf{E})$, where $ch(\mathbf{E})$ is the convex hull of $\mathbf{E}$. By Lemma 3, for each $\nu_k \in \mathbf{M}_\varepsilon$, where $\mathbf{M}_\varepsilon$ is as in Prop. 1, there are at most $N + 1$ points $\underline{f}_{x_1^{\{k\}}}, \cdots, \underline{f}_{x_L^{\{k\}}}$ in $\mathbf{E}$ such that

$$\underline{\nu}_k = \sum_{i=1}^{L} p_i^{\{k\}} \underline{f}_{x_i^{\{k\}}},$$

where

$$p_i^{\{k\}} \geq 0, \sum_{i=1}^{L} p_i^{\{k\}} = 1.$$

Define the probability measures $\{\nu'_k\}$ on $(\mathbf{X}, \mathcal{X})$ by

$$\nu'_k(\{x\}) = \begin{cases} p_i^{\{k\}} & \text{if } x = x_i^{\{k\}}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{A} = \cup_k \left\{ x_1^{\{k\}}, \cdots, x_L^{\{k\}} \right\}$. Clearly, the cardinality of $\mathbf{A}$ is

$$L(\varepsilon) = \|\mathbf{A}\| \leq (N + 1) \times \left\lceil \frac{2\beta}{\varepsilon} \right\rceil^{\|\mathbf{F}\|}.$$

Then the measures ν'_k are as required by the Lemma. □

Appendix B: Proof of Theorem 3

By Lemma 2 and Prop. 1, there is a set of measures with finite support $\mathbf{M}_\varepsilon = \left\{ \nu'_1, \nu'_2, \cdots, \nu'_{N_\varepsilon} \right\}$ such that

$$\sup_{\mu \in \mathbf{M}} \min_{1 \leq i \leq N_\varepsilon} d_{\mathbf{F}}(\mu, \nu'_i) \leq \frac{\varepsilon}{2},$$

where the supremum is over all measures in $\mathbf{M}_{\mathbf{P}_r}$ and

$$N_\varepsilon \leq \left\lceil \frac{4\beta}{\varepsilon} \right\rceil^{\|\mathbf{F}\|},$$

where β is defined as in Prop. 1. Since the mass on each of the supporting atoms can be approximated as closely as desired by rational numbers and the fact that rational numbers are computable in the sense of Section 3.2, then it is easy to see that for each $\nu'_k \in \mathbf{M}_\varepsilon$ there is a computable probability mass function (in the sense of Def. 6) μ_k such that $d_{\mathbf{F}}(\nu'_k, \mu_k) \leq \frac{\varepsilon}{2}$ and, hence,

$$\sup_{\mu \in \mathbf{M}} \min_{1 \leq i \leq N_\varepsilon} d_{\mathbf{F}}(\mu, \mu_i) \leq \varepsilon,$$

where the supremum is over all measures in $\mathbf{M}_{\mathbf{P}_r}$. Note that the measures $\{\mu_i\}$ are not necessarily in $\mathbf{M}_{\mathbf{P}_r}$[3]. We may assume that ε is a computable number w.l.o.g.. Consider the following construction of P:

- Choose any measure $\nu_0 \in \mathbf{M}_{\mathbf{P}_r}$. Let μ_0 be any computable probability mass function such that $d_{\mathbf{F}}(\nu_0, \mu_0) < \varepsilon$.
- Define N_ε counters $i(1), \cdots, i(N_\varepsilon)$ and set them to 0.
- For each $k > 0$ define,
  - $(\forall \psi \in \Psi_0)$, $\bar{\nu}_{\psi,k-1} = \frac{1}{\lambda_{\psi,k-1}} \sum_{l=1}^{k-1} \psi(x^{l-1})\nu_l$ if $\lambda_{\psi,k-1} > 0$, and $\bar{\nu}_{\psi,k-1} = \mu_0$ otherwise.
  - $\alpha_k = 0$ if $(\forall \psi \in \Psi_0)$ $\psi(x^{k-1}) = 0$, and $\alpha_k = \max_{\psi \in \Psi_0} \left\{ d_{\mathbf{F}}(\bar{\nu}_{\psi,k-1}, \mu_0) : \psi(x^{k-1}) = 1 \right\}$ otherwise.
  - $j_k = \operatorname{argmin} i(j)$.

  Note that α_k depends only on x^{k-1}.
- If $\alpha_k > \varepsilon$, let $\nu_k = \mu_0$. Otherwise, let ν_k be the computable probability measure μ_{j_k} and increment $i(j_k)$ by 1.
- Generate x_k according to ν_k.

Note that all the steps in the construction are computable, with the exception of the generation of the outcomes.

Proposition 2. *For $\varepsilon > \beta/m$ and large enough n, P is $\mathbf{F}$-temporally homogeneous $(\Psi, 6\varepsilon, \delta, m, n)$, with*

$$\delta = 2\|\mathbf{F}\|\|\Psi_0\| \mathrm{e}^{-\frac{\varepsilon^2 m^2}{8\beta^2 n}}.$$

[3] Although we think that this restriction can easily be removed, it does not pose any problem to the proof of the theorem.

Proof. Suppose that there is some $\psi \in \Psi_0$ such that $d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \mu_0) > \varepsilon$ and $\lambda_{\psi,n} \geq m$. Let

$$\delta(\mu_0) = \max_{\nu \in \mathbf{M}_\varepsilon} d_{\mathbf{F}}(\mu_0, \nu).$$

Since, by construction, as soon as $d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \mu_0) > \varepsilon$ outcomes start to be generated according to μ_0, then we must have

$$d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \mu_0) < \frac{(\lambda_{\psi,n} - 1)\varepsilon + \delta(\mu_0)}{\lambda_{\psi,n}} \leq \varepsilon + \frac{\beta}{m} \leq 2\varepsilon.$$

Since by Theorem 1 we have

$$P\left(\max_{\psi \in \Psi_0} \{ d_{\mathbf{F}}(\bar{\mu}_{\psi,n}, \bar{\nu}_{\psi,n}) : \lambda_{\psi,n} \geq m \} \geq \varepsilon \right) \leq \leq 2\|\mathbf{F}\|\|\Psi_0\| \mathrm{e}^{-\frac{\varepsilon^2 m^2}{8\beta^2 n}},$$

the proposition is proved. □

Proposition 3. *Let*

$$n \geq \frac{\delta(\mu_0) N_\varepsilon m}{\varepsilon} \|\Psi_0\| + N_\varepsilon m - 1.$$

Then

$$\sum_{j=1}^{N_\varepsilon} i(j) \geq N_\varepsilon m, \tag{1}$$

and, hence,

$$\min_{1 \leq j \leq N_\varepsilon} i(j) \geq m. \tag{2}$$

Proof. We call k an **exceeding time** when

$$\alpha_k = \\ = \max \left\{ d_{\mathbf{F}}(\bar{\nu}_{\psi,k-1}, \mu_0) : \psi \in \Psi_0,\ \psi(x^{k-1}) = 1 \right\} > \\ > \varepsilon.$$

By the construction of P, it is clear that Eqn. 2 follows immediately from Eqn. 1. Suppose that Eqn. 1 does not hold. This means that there have been at least $(n - N_\varepsilon m + 1)$ exceeding times. Since by hypothesis

$$\frac{\delta(\mu_0) N_\varepsilon m}{\varepsilon} \|\Psi_0\| \leq n - N_\varepsilon m + 1,$$

there must be a $\psi \in \Psi_0$ such that, for its corresponding subsequence, $d_{\mathbf{F}}(\bar{\nu}_{\psi,k}, \mu_0)$ has been greater than ε at least $\frac{\delta(\mu_0) N_\varepsilon m}{\varepsilon}$ times. Note that, for each exceeding time

$$\varepsilon < d_{\mathbf{F}}(\bar{\nu}_{\psi,k}, \mu_0) \leq \\ \leq \frac{(\lambda_{\psi,k} - \lambda_{\psi,k,\mathbf{M}}) \times 0 + \lambda_{\psi,k,\mathbf{M}} \delta(\mu_0)}{\lambda_{\psi,k}} = \\ = \frac{\lambda_{\psi,k,\mathbf{M}}}{\lambda_{\psi,k}} \delta(\mu_0),$$

where $\lambda_{\psi,k,\mathbf{M}}$ is the number of times, along the subsequence selected by ψ, such that $\alpha_k \leq \varepsilon$. From the last inequality, it follows that

$$\lambda_{\psi,k,\mathbf{M}} > \frac{\varepsilon}{\delta(\mu_0)}\lambda_{\psi,k}.$$

Therefore, for ψ's last exceeding time we have

$$\lambda_{\psi,k,\mathbf{M}} > \frac{\varepsilon}{\delta(\mu_0)}\lambda_{\psi,k} \geq \frac{\varepsilon}{\delta(\mu_0)}\frac{\delta(\mu_0)N_\varepsilon m}{\varepsilon} = N_\varepsilon m.$$

However, this contradicts our initial assumption that there were less than $N_\varepsilon m$ exceeding times along the entire sequence. Thus, we must conclude that Eqn. 1 holds. □

Let $\Psi_1 = \{\psi_1, \psi_2, \cdots, \psi_{N_\varepsilon}\}$ be a set of N_ε place selection rules such that, for $1 \leq l \leq N_\varepsilon$, ψ_l selects the subsequence where the measure μ_l has been used. The fact that such a family Ψ_1 of computable place selection rules exists follows from the construction of P. Note that Proposition 3 implies that the subsequences selected by the rules in Ψ_1 have length larger than or equal to m.

Proposition 4. $\mathbf{M}$ *is* $\mathbf{F}$*-visible* $(\Psi_1, 3\varepsilon, \delta, m, n)$*, where*

$$\delta = 2\|\mathbf{F}\|\|\Psi_1\|\mathrm{e}^{-\frac{\varepsilon^2 m^2}{8\beta^2 n}}.$$

Proof. It is clear that, by construction, for all $\mu \in \mathbf{M}$ there is a measure $\mu_i \in \mathbf{M}_\varepsilon$ and a rule $\psi \in \Psi_1$ such that

$$d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \mu) \leq d_{\mathbf{F}}(\bar{\nu}_{\psi,n}, \mu_i) + d_{\mathbf{F}}(\mu_i, \mu) \leq \varepsilon + \varepsilon \leq 2\varepsilon.$$

Then the proposition follows from Theorem 1 and the fact that Proposition 3 implies that $(\forall \psi \in \Psi_1)\ \lambda_{\psi,n} \geq m$. □

The **proof of Theorem 3** follows from Propositions 2-4.

References

[1] Mark Braverman. On the complexity of real functions. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*. Technical Committee on Mathematical Foundations of Computing, IEEE Computer Society, 2005.

[2] Mark Braverman and Stephen Cook. Computing over the reals: Foundations for scientific computing. *Notices of the AMS*, 53(3):318–329, March 2006.

[3] Abbas Edalat. Domains for computation in mathematics,physics and exact real arithmetic. *The Bulletin of Symbolic Logic, Association for Symbolic Logic*, 3(4):401–452, December 1997.

[4] Pablo I. Fierens. *Towards a Chaotic Probability Model for Frequentist Probability*. PhD thesis, Cornell University, 2003.

[5] Pablo I. Fierens and Terrence L. Fine. Towards a frequentist interpretation of sets of measures. In G. de Cooman, T. L. Fine, and T. Seidenfeld, editors, *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*. Shaker Publishing, 2001.

[6] Pablo I. Fierens and Terrence L. Fine. Towards a chaotic probability model for frequentist probability. In J.M. Bernard, T. Seidenfeld, and M. Zaffalon, editors, *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*. Carleton Scientific, 2003.

[7] Peter Hertling. Computable analysis via representations. Notes for a presentation made at a Satellite Seminar of the Second International Conference on Computability and Complexity in Analysis, held in Kyoto, Japan, August 2005.

[8] Leandro Chaves Rêgo and Terrence L. Fine. Estimation of chaotic probabilities. In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, 2005.

[9] Walter Rudin. *Functional Analysis*. McGraw-Hill, 1973.

[10] Glenn Shafer and Vladimir Vovk. *Probability and Finance. It's Only a Game!* Wiley Series in Probability and Statistics. John Wiley & Sons, 2001.

[11] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.

[12] Klaus Weihrauch. A simple introduction to computable analysis (2nd edition). Informatik Berichte 171, FernUniversitt Hagen, Hagen, July 1995.